DOCUMENT RESUME

ED 060 876 LI 003 542

AUTHOR	Roberts, Stephen A.
TITLE	A Machine Readable Data Base of Social Science Serials.
INSTITUTION	Bath Univ. of Technology (England). Univ. Library.
PUB DATE	Nov 71
NOTE	27p.; (7 References); Design of Information Systems in the Social Sciences, Working Paper 2
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Data Bases; *Design; Electronic Data Processing; Foreign Countries; *Information Systems; *Serials; *Social Sciences
IDENTIFIERS	England; *Machine Readable Data

ABSTRACT

A discussion of one of the main activities of the DISISS (Design of Information Systems in the Social Sciences) project is presented in this working paper. A comprehensive Check List of Social Science Serials (CLOSSS) is being prepared as a basis for bibliometric studies of the primary and secondary literature of the social sciences. The CLOSSS data base will be made machine readable so that a wide range of bibliometric studies can be undertaken using statistical and mathematical techniques. Section 2 of this paper deals with requirements for bibliometric studies. Problems associated with the construction of a machine readable data base are discussed in section 3. Detailed design features and problems of CLOSSS are set down in section 4. Section 5 takes a wider view and looks at CLOSSS in a number of contexts in which a serials data base can make a contribution. Section 6 concludes the discussion giving some general issues which must be resolved before the future development of CLOSSS can be estimated. (Author/NH)

ED 060876

Bath University Library

Design of Information Systems in the Social Sciences,

Working Paper No. 2

A MACHINE READABLE DATA BASE OF SOCIAL SCIENCE SERIALS

Stephen A. Roberts

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

November 1971

LI 003 542

CONTENTS

		Page
1.0	INTRODUCTION	1
2.0	REQUIREMENTS OF BIBLIOMETRIC STUDIES	1
	2.1 Basic pattern of studies	1
	2.2 Social science serials data	2
3.0	CONSTRUCTION OF A MACHINE READABLE DATA BASE	4
	3.1 Alternative serials systems and record formats	4
	3.2 Current situation	6
4.0	DESIGN FEATURES AND TECHNICAL SPECIFICATIONS	8
	4.1 Data elements and record formats	8
	4.2 Data collection for CLOSSS	8
	4.3 Editing, coding and data input	11
	4.4 Programs	11
	4.5 DISISS data bases	11
5.0	USES OF MACHINE READABLE DATA BASES WITH PARTICULAR REFERENCE TO CLOSSS	12
	5.1 National and international serials data projects	12
	5.2 Bibliographic reference tools	15
	5.3 Union listing	15
	5.4 Updating the UNESCO World List of Social Science Periodicals (WLOSSP)	15
	5.5 Library management	16
	5.6 Bibliometric applications	17
6.0	CONCLUSION	18
References		20
Table 1	Data elements for CLOSSS	7
Table 2	Comparative coverage of data elements	9
Diagram 1	Exchange of serials data	13
Appendix:	Data collection sheets Pilot version Final version	

1.0 INTRODUCTION

The purpose of this working paper is to present a discussion of one of the main activities of the DISISS (Design of Information Systems in the Social sciences) project based at the University of Bath under the direction of Maurice B. Line. The project runs from January 1971 to December 1973 and is supported by a research grant from OSTI.

A comprehensive Check List of Social Science Serials (CLOSSS) is being prepared as a basis for bibliometric studies of the primary and secondary literature of the social sciences. The CLOSSS data base will be made in machine readable form so that a wide range of bibliometric studies can be undertaken using statistical and mathematical techniques.

The various issues involved in the production of CLOSSS are presented at some length owing to the complexity of the task itself and to the potential interest and value that the undertaking has for workers in information, documentation and bibliography.

Section 2.0 deals with requirements for bibliometric studies, and is followed by a discussion of the need for a machine readable data base. Problems associated with its construction are discussed, and a résumé of progress to date is given in section 3.0.

Detailed design features and problems of CLOSSS are set down in section 4.0. Section 5.0 takes a wider view and looks at CLOSSS in a number of contexts in which a serials data base can make a contribution. Section 6.0 concludes the discussion.

2.0 REQUIREMENTS OF BIBLIOMETRIC STUDIES

2.1 Basic pattern of studies

A bibliometric study of the primary and secondary literature of the social sciences requires the construction of a data base of serials (CLOSSS) for two main purposes.

(a) A descriptive bibliometric survey of the formal system. At present very little is known about the size, growth, mortality, linguistic and geographical composition, subject scattering and density of the serial and monograph literature of the social sciences. DISISS will cover the serial literature in considerable depth; due to problems of data collection and sampling, the monograph literature will receive less detailed study. A sound description of the formal system is a prerequisite of any information systems design.

(b) Large scale citation studies of the social science literature are being undertaken to investigate structural characteristics of the formal system. Citations from samples of social science sources are to be analysed by statistical and mathematical methods, in particular cluster analysis. Journal clusters and author networks will be developed empirically from citation data and compared with existing subject groupings in documentation and social science. Coverage patterns derived from journal clusters will be used in computer based modelling of secondary services and as criteria for coverage in systems designs.

For these analytical studies a data base of serials is required for the identification and development of matrices on which citations can be recorded and analysed.

These varied studies, involving the handling of very large amounts of data, e.g. 10,000-15,000 serial records and at least 50,000 citation records, can best be accomplished by using computerized data bases. CLOSSS represents the primary data base; citations form another machine readable data base. Machine readable data will also be employed in some studies of coverage of abstracting and indexing services. Modelling of system components using bibliometric, user and cost information will also require use of computerized data bases.

2.2 Social science serials data

The following specifications are being observed for the preparation of the machine readable data base of social science serials. No existing data base meets the criteria listed here.

(a) Coverage. The descriptive parametric studies and the citation studies require the most comprehensive coverage of serial titles possible. CLOSSS will attempt to record data on all serial titles relevant to the social sciences.

(b) Statistical requirements. The citation studies require extensive use of sampling techniques if the results are to be at all valid. In the past citation studies have had a poor reputation for statistical validity (Raisig, 1962) and it has not been possible to draw firm conclusions from the results of the studies. A comprehensive CLOSSS can provide a high quality sampling frame for citation studies of the social sciences, particularly the sampling of source journals.

(c) Data elements. The selection of data elements made for CLOSSS is designed to be of maximum value for descriptive bibliometric studies (Table 1). Existing bibliographical sources of serials data, of which none are generally available or are in machine readable form, do not give the coverage or quality of data required.

(d) Quality and control of data. For the descriptive analysis it is desirable that full data should be collected for all serials recorded on CLOSSS. The information must be as accurate as possible. Data collection requires a considerable level of control. Where possible each title is examined in the field so that published bibliographical data can be checked and supplemented before input to the CLOSSS file.

(e) Research design. To give flexibility to the research programme it is desirable to have as complete an integration as possible of all stages of the work from data collection and storage, to processing and analysis. A purpose designed system for CLOSSS and the other related data bases is the best way to achieve consistency and compatibility among available data.

There is currently considerable interest in machine readable serials files relevant to the social sciences. Projects exist in France (Groupe de Travail sur l'Analyse des Périodiques; UNESCO/World List of Social Science Periodicals; International Serials Data System), in the United Kingdom (development of serial files for the British Library) and in the United States of America (US National Libraries

Task Force; Bowker & Co. Ltd.). The work ranges from applications in library housekeeping to information storage and retrieval. Considerable discussion centres around the record formats to be adopted for the machine readable data. The MARC formats regarded as suitable for the international exchange of all types of bibliographic data, have received attention from the serials projects. Recently an abbreviated MARC format for serials, entitled MASS (Marc-based Automated Serials System) has been developed and is being considered by SCONUL[1] in the UK. MASS is fully compatible with MARC.

A comprehensive machine readable file of social science serials data could be an appropriate contribution to the various current and probable future efforts. CLOSSS has a potential beyond its use as a research data base. The possibilities of compatible files, exchange of data and coordination of social science serials information have been examined by the DISISS team.

The CLOSSS file is to be created in a form compatible with the MASS data requirements, but at present using simplified local tags and a local machine format. Creation of the file will take place between late 1971 and early 1972.

3.0 CONSTRUCTION OF A MACHINE READABLE DATA BASE

In the short term, the bibliometric research work of the project will assemble valuable empirical data for exploratory studies in the design of information systems. In the longer run it is desirable and feasible to consider production of the data base in a form most usable by and relevant to other projects concerned with serials data in general, and bibliometrics in particular. The long term considerations have been responsible for our development of a record format eventually compatible with MARC and MASS. The international exchange of bibliographic data will become increasingly important in the near future, and the standardization of record formats is vital.

3.1 Alternative serials systems and record formats

The decision to develop CLOSSS with MASS compatible data

---

[1] SCONUL: Standing Conference of National and University Libraries.

was not made without consideration of a number of alternative approaches. Four other systems have been considered.

(a) Loughborough Periodicals Data System. When use of the MASS format was first considered, the DISISS project was unable to cope with the technical problems which would have arisen in attempting to use unmodified the existing MASS record structure and package programs (as then developed) on the ICL System 4 computer available at Bath. The Loughborough PDS system appeared to offer an attractive compromise, since in the long term it was likely to be related to a MASS based system. Discussions with Loughborough proved fruitful, and with some modifications CLOSSS could have been produced in PDS on an ICL 1907 computer or, with additional work on conversion routines, on an ICL System 4 computer (Loughborough University of Technology, 1971). This option has been discarded since DISISS has obtained the services of a computer programmer for the project.

(b) Full MARC records. Full MARC records (Library of Congress, 1969), despite growing international favour especially in the USA and in libraries, would prove cumbersome if applied to CLOSSS. MARC serial records are lengthy, occupy a large amount of computer store and would require a great deal of effort in the later stages of statistical and mathematical analysis of data files. In any case, MASS represents a very useful subset of the full MARC record.

(c) ISDS format. The International Serials Data System (ISDS) proposal (Martin & Barnes, 1970) has a number of attractions, particularly as a scheme of international and comprehensive scope. The project is related to the use of ISSNs and is connected with the overall developments of bibliographic description standards envisaged by the International Standards Organization and UNISIST. At the present time there is no final agreement on the format, although it would probably not be dissimilar to MASS (see Table 2) even if some conversion of tags was necessary to achieve compatibility.

(d) MASS. The MASS system has been developed at Loughborough and Birmingham University Libraries. It is fully MARC compatible although it does not use all the MARC fields. MASS is a logical adaptation of MARC well suited to local needs. The possibility of opportunities for local use

and its overall compatibility with the MARC format make it seem ideal for CLOSSS. There have been discussions between DISISS and MASS projects and although the obstacles of different field specification were resolved, there still remained the problems of programs and computer facilities. No MASS package programs were likely to be available in 1971 or early 1972, and those that were could not be run on ICL System 4 computers.

(e) Local MASS compatible system. A locally developed machine readable data base system using a local record format was given a pilot trial in January 1971. The system which is now considered most suitable for CLOSSS is essentially local, although potentially MASS compatible. This approach makes the best use of programming and computing resources available at Bath, and fits in well with program and system development for cluster analysis and modelling. A subset of the MASS data requirements has been chosen for CLOSSS.

3.2 Current situation

Creation of the machine readable CLOSSS file is to take place from November 1971, and should take about 6 months at least, using information on serial titles collected during field exercises in April, July, September and November 1971.

Collection of serials data has gone ahead rapidly, in spite of difficulties in reaching a suitable decision on record format and computer facilities. By October 1971 information had been gathered on 5,500 serial titles, and about 7,000 fully edited records will be on the file early in 1972. It is hoped to continue data collection, subject to time and cost constraints, during the period to mid-1973. Information on new serials derived from the machine readable file at the NLLST could be used with CLOSSS and this possibility is being explored by DISISS. Serial titles revealed by citation studies will also be added to the CLOSSS file.

Consultations with individuals and groups interested in serials data has continued during 1971, and in September a letter outlining the DISISS project and the CLOSSS proposal was circulated to a selection of experts in the UK, USA and Europe.

TABLE 1

DATA ELEMENTS FOR CLOSSS

(00) Identification code

(01) Title (in full, as it appears currently)

(02) Title in English if different from (01)

(03) Alternative title(s)

(04) Previous title(s)

(05) Beginning date

(06) Ending date, if any

(07) Frequency/issues per annum

(08) Issuing body

(09) Type of issuing body

(10) Publisher (Name)

(11) Country of publication

(12) Type of serial

(13) Description of serial

(14) Nature of contents

(15) Abstracts with articles

(16) Language(s) of contents

(17) Assessment of subject content

(18) Number of articles in 1969

(19) Subscription price (1969)

(20) Coverage by indexing and abstracting services

(21) ISSN

4.0 DESIGN FEATURES AND TECHNICAL SPECIFICATIONS

4.1 Data elements and record formats

The data elements selected for CLOSSS (Table 1) have two basic functions. Firstly, as descriptive cataloguing data, e.g. title, previous title, issuing body; secondly as bibliometric data, which although strictly of a bibliographic nature is considered more important as descriptive parametric information, e.g. beginning date, frequency in issues per year, country of publication, subject content. The variables in the bibliometric analysis consist of the second type of data. However, for file creation purposes there is no difference between the two types of data. The definition of data elements for CLOSSS owes much to the standards established by Loughborough PDS (Evans, Wall & MacKay, 1969) and the MASS proposal (Loughborough University of Technology and Birmingham Libraries Co-operative Mechanization Project, 1970).

Table 2 shows the relationship between CLOSSS and the serials systems considered by DISISS, in terms of record fields and format. Both MASS and a MASS minimal subset can accommodate all or a major selection of CLOSSS data elements in their record structure. This correspondence of data elements makes it possible for CLOSSS to use the MASS tag structure if conversion routines are developed; adopting the compatible tag structure provides the key to any future exchange of serials data between CLOSSS and files using MARC or MASS record formats. The acceptance of a standard record format and set of tags is of great importance to the current international effort on the control of serials data and the creation of machine readable files. In addition there must be flexibility to allow insertion of locally required data elements. Particularly important is the comparison in Table 2 of data elements for CLOSSS and ISDS. The overlap would enable CLOSSS to make initially a substantial contribution to an international file of serials data. Data elements required for DISISS studies of the secondary serial literature are shown; they form a subset of CLOSSS data elements. To accord with future international practice ISSN has been added to the set of data elements for CLOSSS, although it was not included on the data collection sheets.

4.2 Data collection for CLOSSS

Specimens of data collection sheets are appended to this paper. The earlier version was used in April 1971 and subsequently modified to provide

TABLE 2

COMPARATIVE COVERAGE OF DATA ELEMENTS

CLOSSS (Collection sheet)	LUT/PDS	MASS MINIMAL SUBSET	ISDS (MARTIN & BARNES) (Levels 0-3)	FULL MASS	SECONDARY SERVICES DATA : REQUIREMENTS
(00) IDENTIFICATION NUMBER	X	X G/001	X D3	X G/001	X
(01) TITLE	X	X G/200 (General variable)	X D11/D14(?)	X G/200	X
(02) TITLE IN ENGLISH IF DIFFERENT FROM (1) (03) ALTERNATIVE TITLE	X	X G/246 (General variable)	X D13	X G246	X
(04) PREVIOUS TITLE	X	X G/780 (General variable)	X D17	X G/247	X
(05) BEGINNING DATE	X	X G/008C	X D6	X G/008C	X
(06) ENDING DATE	X	X G/008D	X D7	X G/008D	X
(07) FREQUENCY ISSUES PER YEAR	X		X D8	X G/008F	X
(08) ISSUING BODY (09) TYPE OF ISSUING BODY	X	G/260(?) (General variable)	X D32	X G/260 (ǂe)	X
(10) PUBLISHER (NAME)	X	X G/260 (General variable)	X D24	X G/260 (ǂb)	X
(11) COUNTRY OF PUBLICATION	X	X G/008E	X D25	X G/008E	X
(12) TYPE OF SERIAL	X	X G/008I	X D33	X G/008I	X
(13) DESCRIPTION OF SERIAL	X	X G/008L		X G/008L	
(14) NATURE OF CONTENTS	X			X G/008M	
(15) ABSTRACTS WITH ARTICLES	X			Temporary additional field	
(16) LANGUAGE OF CONTENTS	X	X G/008U (G/041)	X D10	X G/008U	X
(17) SUBJECT CONTENT	X	Temporary additional field	X D26	Temporary additional field	X
(18) NUMBER OF ARTICLES	X			Temporary additional field	X
(19) SUBSCRIPTION PRICE	X			X G/350	X
(20) INDEXING SERVICES (COVERAGE)			X D28	X G/510	
(21) ISSN		G/001	X D3	X G/022	X

the version used from July 1971, which is now the standard data collection sheet for serials data. The data collection sheets went through extensive field trials at the Polytechnic of North London School of Librarianship. Research assistants and students at the Polytechnic have been responsible for the bulk of field data collection at Senate House Library, University of London; British Library of Political and Economic Science; National Lending Library for Science and Technology; and a variety of specialist subject libraries, mostly in the London area.

Bulk collection of data in the field is supplemented by the identification of serial titles in published bibliographies. Serial titles discovered in this way are sought in libraries (where possible) and information on them collected by examination of actual copy.

The main published bibliographic sources in use are listed below.[1] Each source contributes at least a small proportion of new titles. Collection policy has been to err on the side of exhaustivity and inclusiveness with regard to subject coverage and type of publication. The citation data should reveal a number of new and relevant titles to add to the file, particularly from non-conventional social science fields. This process of addition constitutes the bibliometric refinement of the serials file. At some stage certain titles might be removed from the file. Although data collection has drawn on the resources of many libraries, the overall rate of duplication on titles already collected is so far about 6 per cent.

---

[1] British union catalogue of periodicals (BUCOP) London, Butterworths, 1955-1958; Inventaire des périodiques étrangers et des publications en série étrangères reçus en France, Paris, Bibliothèque Nationale, 1969; Irregular serials and annuals: an international directory, New York, Bowker, 1967; National lending library for Science and Technology, Current serials received, London, HMSO, 1971 (an up-to-date printout was made available to DISISS); New periodical titles, London, Butterworths, 1964- ; Ulrich's international periodicals directory, New York, Bowker, 1970; World list of social science periodicals, Paris, UNESCO, 1966.

In addition a file of specialist bibliographic sources of information on serials is maintained to supplement the general works mentioned above.

4.3 Editing, coding and data input

The Polytechnic of North London has been responsible for full editing of data. Published sources, staff expertise and further field collection have assisted in improving the quality and coverage of data. Although great use has been made of published sources initially to locate titles, their use in editing has been less. Each source has certain advantages and disadvantages for the purpose of collecting and editing serials data and this lends weight to our initial premise that no existing list of social science or other serials was directly suitable for DISISS bibliometric studies.

The data is to be punched on to paper tape and a master file created on magnetic tape or disk. Paper tape input is more suitable to the variable length data constituting the CLOSSS records. Some data elements are being replaced by codes on the data collection sheet at the final editing stage and punching will be done directly from edited sheets. It has been possible to dispense with a manual coding from data collection sheets to a computer coding form.

Coding and punching errors will be identified, either by computer during the data vetting or by visual proof reading when the proofs of the file print-out are examined.

4.4 Programs

Programs for the data vetting, descriptive statistical analysis of the file and setting up of matrices for the citation analysis will be developed on the project. Some use may be made of standard statistical packages. Later on, conversion programs for MASS or a similar internationally approved record structure could be developed.

4.5 DISISS data bases

As far as possible internal consistency will be sought between serials, citations, secondary services and modelling data bases. Attention is being given to the use of other files of data which could be useful for bibliometric analysis. These other files may or may not yet be compatible with internal DISISS files. Psychological Abstracts

tapes, the ISI Science Citation Index file, and the Bowker machine readable serials file are potentially useful.

If consistency was achieved through MASS compatible record formats, it would be possible to run bibliometric analyses on other data using DISISS programs. The type of products and exchange of information outlined in Chapter 5 of the ISDS proposal (Martin & Barnes, 1970) could be realised, e.g. periodic bibliometric review projects (para. 5.4.3). There are good reasons for this broader outlook - data collection can be reduced; other sources could be used to widen the sample for bibliometric studies; bibliometric techniques could be tested in other situations; and some return on program development effort could be expected.

There appears to be every advantage to be gained from creating compatible machine readable data files.

5.0 USES OF MACHINE READABLE DATA BASES WITH PARTICULAR REFERENCE TO CLOSSS

When CLOSSS becomes available in MASS or an internationally compatible format there are a number of areas where it would be desirable and feasible to take advantage of its compatible machine readable structure.

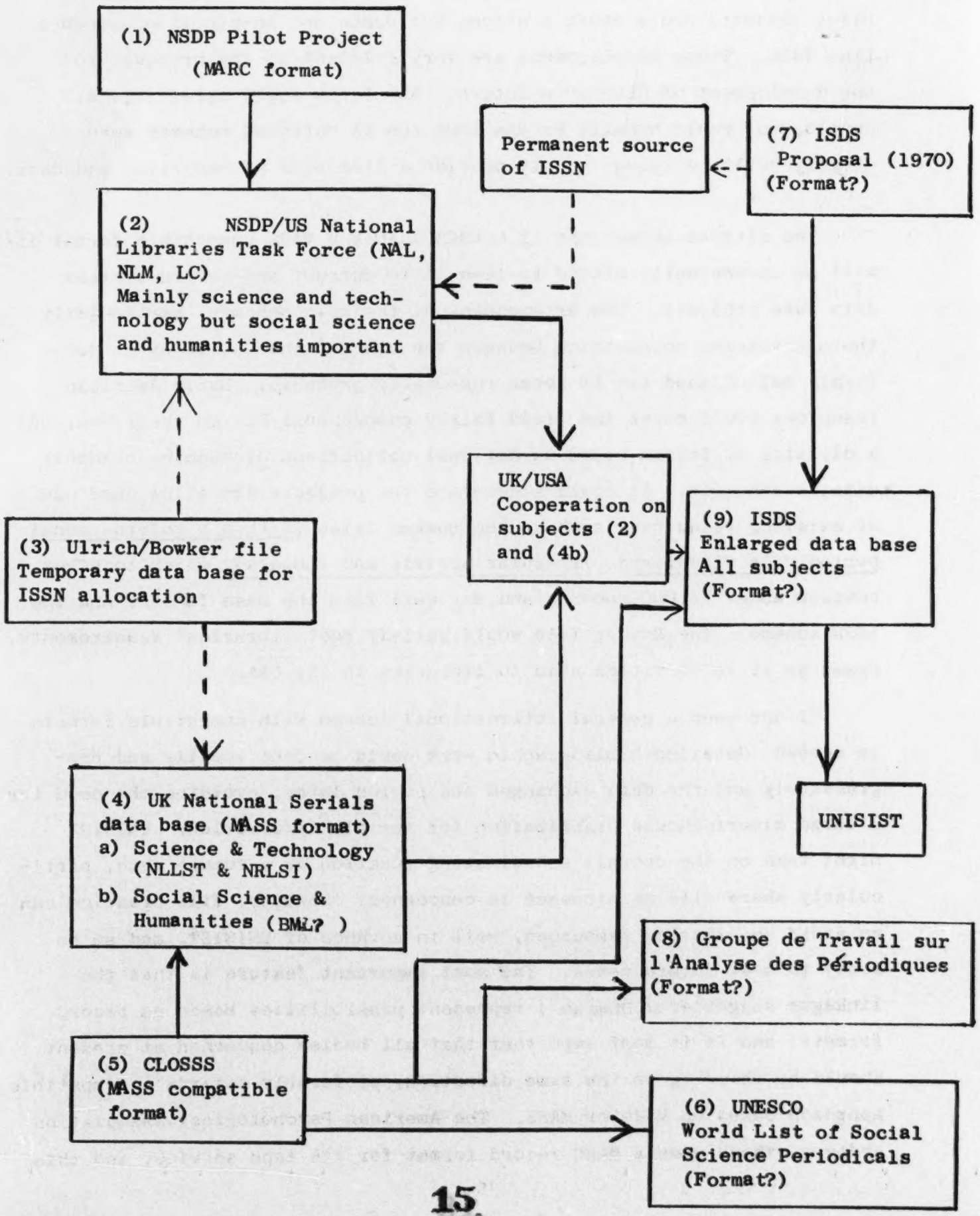
5.1 National and international serials data projects

Diagram 1 sets out some of the features of the present international situation on serials mechanization. In most cases the ultimate objectives of the projects are the creation and maintenance of comprehensive files of machine readable serials data which can be exploited in a variety of ways.

The relationships in Diagram 1 are shown through format specifications for the serials records. There is no coordinated plan as yet developed for all the various schemes, but potential opportunities for closer links are apparent. It would appear not unreasonable to press for the adoption of one standard record format and tagging structure based on MARC/MASS proposals, irrespective of number and type of data elements. Present international thinking seems to be moving, however loosely, towards coordination. A catalyst is required at this stage in the form of a firm commitment to a MARC/MASS format. CLOSSS could provide such

DIAGRAM 1

EXCHANGE OF SERIALS DATA


(1) NSDP Pilot Project (MARC format)
Permanent source of ISSN
(7) ISDS Proposal (1970) (Format?)
(2) NSDP/US National Libraries Task Force (NAL, NLM, LC) Mainly science and technology but social science and humanities important
UK/USA Cooperation on subjects (2) and (4b)
(9) ISDS Enlarged data base All subjects (Format?)
(3) Ulrich/Bowker file Temporary data base for ISSN allocation
(4) UK National Serials data base (MASS format)
a) Science & Technology (NLLST & NRLSI)
b) Social Science & Humanities (BML?)
UNISIST
(8) Groupe de Travail sur l'Analyse des Périodiques (Format?)
(5) CLOSSS (MASS compatible format)
(6) UNESCO World List of Social Science Periodicals (Format?)

a stimulus if it goes ahead as planned, since it will represent within 12-18 months a uniquely comprehensive list of social science serials meeting a significant portion of MARC/MASS data requirements.

For national and international schemes CLOSSS could form the source of social science serials records. Developments in the next 12-18 months in UK and USA national projects (e.g. British Library, National Serials Pilot Project) could exert a strong influence on 'in-pipeline' schemes like ISDS. These developments are very relevant to any proposal for the development of CLOSSS in future. Any large scale international development would benefit in the long run if national schemes were clearly outlined soon to provide a firm base of expertise and data.

The diagram shows that if CLOSSS adopts a MASS compatible format it will be conveniently placed to feed in to current and future serials data base projects. The arrangement of the main schemes, particularly those involving cooperation between the USA and the UK, is by no means firmly established but it seems reasonably probable. North American resources could cover the field fairly comprehensively on their own, but a division of labour based on national collections of machine readable data is sensible. It could accelerate the projects and allow good use of existing resources, such as the Bowker files (Ulrich's international periodicals directory; Irregular serials and annuals), which together contain about 75,000 records and may well form the base for ISDS and the ISSN scheme. The Bowker file would satisfy most libraries' requirements, based as it is on titles sold to libraries in the USA.

If and when a general international scheme with compatible formats is agreed, detailed bibliographic work could be done locally and progressively and the data exchanged and pooled later, avoiding the need for a large clearinghouse organization for serials information. UNISIST might take on the overall coordinating function at a future date, particularly where file maintenance is concerned. However, file creation can go ahead on existing resources, well in advance of UNISIST, and so be ready to meet future needs. The most important feature is that the linkages suggested in Diagram 1 represent possibilities based on record formats, and it is most important that all bodies concerned at present should be thinking in the same direction, preferably towards a compatible approach based on MARC or MASS. The American Psychological Association amongst others uses a MARC record format for its tape service, and this

trend could well be encouraged. Through the use of a common record format a secondary service data base, such as those developed at APA, could be directly fed in to a modelling exercise.

5.2 Bibliographic reference tools

With a machine readable data base it is relatively easy to produce hard copy listings, and the use of computer typesetting would enable a very attractive bibliographic tool to be produced from CLOSSS, if the records were properly tagged. Bibliometric refinements to the data base would make it possible to structure the list in a variety of ways. KWIC and KWOC indexes could be produced and would make the list a useful tool - a feature lacking from many existing lists. On-line searching and selective analyses of the file for bibliographic purposes would be possible. A marketable magnetic tape version could be produced, and such a tape could, if formats allowed, be used with other data bases; there would be no necessary restriction to social science materials.

To achieve these objectives it would be necessary to edit and supplement the present CLOSSS file. DISISS is already considering ways of updating the CLOSSS file. There would not necessarily be a great overlap with ISDS, which at least in the short term does not plan to cover the social sciences. Where possible compatible schemes should be regarded as complementary.

5.3 Union listing

Although computer based systems make union lists easier to maintain, the case for producing them in future could be limited, particularly with the possible development of comprehensive loan collections of serials. A machine readable file such as CLOSSS could potentially contribute to union list projects.

5.4 Updating the UNESCO World List of Social Science Periodicals (WLOSSP)

This list of social science serials was not adequate by itself as a basis for the DISISS bibliometric studies, since it contains only a nucleus of specifically social science serials; in all 1,313 titles are recorded. Although WLOSSP is updated by supplementary information in the International Social Science Journal there is at present considerable delay in disseminating it in a cumulated form.

If CLOSSS reached hard copy in the next two years, when a new edition of WLOSSP is likely to be available, there would be some redundancy. However, the two projects are in some ways of different design and conception, so that the problem of conflict is not so pressing. CLOSSS serials data would certainly be relevant to the maintenance of the WLOSSP file, now being considered for machine readable production.

The whole issue of updating local and pre-existing files is complicated by current uncertainty on international developments. UNISIST, for example, has discussed a central registry of scientific serials, presumably based on ISDS. An international archive of serials data could influence the scope and pace of local developments. Some attempt at least should be made to reduce duplication of effort.

5.5 Library management

Research on quantitative analytical techniques in library management like that done at Lancaster University Library (Buckland et al., 1970), and the analytical bibliometric methods being explored by DISISS on characteristics of secondary services, reviews and citation patterns are both concerned with statistical hypotheses like the Bradford-Zipf-Mandelbrot empirical hyperbolic distributions (Fairthorne, 1969), used to describe features of the documentation and information process. Both areas of research seek to improve efficiency, to provide users with a more effective service and to effect rationalization in the design and operation of information systems.

At some point the applications of statistical hypotheses in library service provision and information/documentation organization may meet to resolve both practical and theoretical problems in information science. Library management applications are currently focussed on the conduct of technical services and on stock control, e.g. loan policies, weeding and acquisition policy.

It is not yet clear how CLOSSS could contribute directly to these areas, but in the form of a comprehensive store of social science serials data, possibly applied to bibliometric analyses of literature, it could be exploited to serve library management ends.

5.6 Bibliometric applications

The techniques of bibliometric investigation and control of information systems are not very advanced at present, nor are they likely to be so for several years. Ideas in the area are still 'soft', compared with 'hard' proposals on data exchange and record formats.

CLOSSS and the other data bases mentioned here form the basis of the DISISS investigations into the applications of bibliometric methods. A major effort is being made towards the development of clustering techniques for use on citation data. Empirically generated journal clusters can be compared with existing subject concepts and the treatment of bibliographic information in subject oriented secondary services, review literature and other information retrieval devices. An extension of the work already undertaken by Salton, Kessler and Garfield based on computerized retrieval systems is a possibility. DISISS will, however, attempt to apply techniques and findings to conventional bibliographic tools and services, rather than to on-line interactive retrieval systems.

Extensive data base development is essential for this experimental work. Moreover the basis of the work can be broadened by access to compatible data. Producing CLOSSS and other data bases with MARC/MASS compatible records seems to be a good way of helping to advance bibliometric work (provided that MARC or similar records are generally accepted), enabling models to be tested in a variety of experimental situations.

Future information systems require predictive and regulatory mechanisms if they are to be sensitive to user requirements. Bibliometric methods could provide the type of control necessary in future systems. Eventually a large proportion of available bibliographic information might be in machine readable form, and the experimental data bases might constitute the whole population or random samples from it. The ISDS proposal points out the analytical potential of a serials data base; DISISS could in effect pilot many of the possible analyses which could then be applied in the context of international information control when a comprehensive system like ISDS becomes operational.

Although, to be realistic, an information control system using computer based bibliometric techniques is going to be at least five years from realization, at a feasibility study level contacts are being made by DISISS with bodies like the Institute of Scientific Information and ICSU/AB. What is essential, and can be done now, is to ensure provision for the exchange of experimental bibliometric and conventional bibliographic data.

6.0 CONCLUSION

From technical and conceptual viewpoints the proposed machine readable data base of social science serials in exchangeable format looks an attractive proposition as a short term research tool, with inherent long term possibilities. Some general issues need to be resolved before the future development needs of CLOSSS can be estimated. Briefly, the major issues are as follows, and relate not only to social science serials:

(a) Record format. This topic has dominated the present discussion and is absolutely central to any international serials data programme. Formats derived from MARC and compatible with it should be considered essential for serials data base projects.

(b) Time schedules. By 1973 the DISISS project will have produced a substantial amount of bibliometric data analyses and software. National and international schemes should have progressed further, and some idea of the situation in about three years' time is required for planning purposes.

(c) Organization and management. The best means of coordinating work on serials data should be sought. The most effective means means of cooperation could be under the auspices of ISDS.

(d) Coverage. A high level of subject coverage of titles, both current and retrospective, is required. The exact division of responsibility between subjects needs investigating in the context of international cooperation. Ultimately a high standard of collection of serials information is required. A standard set of descriptive data elements is needed.

Adequate facilities for data collection are desirable and the planned role for ISDS is an important consideration.

(e) Resources and costs. Reducing the duplication of effort, e.g. through coverage plans, is a positive step towards the best allocation of resources and control of costs. Continued work on CLOSSS as a major file of social science material would require some additional expenditure, although it is unlikely to be excessive.

(f) Strategy for bibliographical control. Such a strategy is implicit in the development of UNISIST and ISDS. It is essential that their scope is enlarged to include the social sciences.

In the immediate future, work on CLOSSS must be more or less self-contained within the DISISS project. However, experience gained and data collected during the project should not be wasted if plans are laid which would facilitate re-use of data. Compatible record formats would enable re-use of data and further allow use of files of data generated by other projects.

CLOSSS is a contribution to a comprehensive programme of bibliometric research, which is potentially useful in assisting the design and operation of information systems. It is essential that such a programme should combine the theoretical with the practical. The type of serials data base proposed in CLOSSS offers this prospect of blending theoretical and practical areas together. On the one hand it promises to be a useful bibliographical tool, on the other a means of achieving certain ends of theoretical research as a result of which practical action can be taken. Central to the realization of theoretical and practical ends is a compatible machine readable serials data base.

REFERENCES

BUCKLAND, M.K., HINDLE, A., MACKENZIE, A.G. and WOODBURN, I. (1970). Systems analysis of a university library: final report on a research project. Lancaster, University of Lancaster Library, 1970. (University of Lancaster Library Occasional Papers, No. 4). Particularly Chapter 3, The structure of library usage.

EVANS, A.J., WALL, R.A. and MACKAY, J.C. (1969). Periodicals data automation project. Loughborough, Loughborough University of Technology, 1969.

LIBRARY OF CONGRESS, INFORMATION SYSTEMS OFFICE, (1969). Serials, a MARC format: working document. Washington, Library of Congress, 1969.

LOUGHBOROUGH UNIVERSITY OF TECHNOLOGY (1971). Extension project on library automation: progress April to July, 1971. Loughborough, Loughborough University of Technology, 1971. (Mimeo).

LOUGHBOROUGH UNIVERSITY OF TECHNOLOGY LIBRARY and BIRMINGHAM LIBRARIES CO-OPERATIVE MECHANIZATION PROJECT (1970). MASS (MARC based Automated Serials System) Working paper No. 1. Loughborough, Loughborough University of Technology, 1970.

MARTIN, M.D. and BARNES, C.I. (1970). Report on the feasibility of an International Serials Data System, prepared for UNISIST/ICSU - AB Working Group on Bibliographic Descriptions. London, INSPEC/IEE, 1970, DM/CB/284.

RAISIG, L.M. (1962). Statistical bibliography in the health sciences. Bull. Med. Lib. Assn. 50(3), July 1962, 450-461.

APPENDIX

Data Collection Sheets

PILOT VERSION

DISISS - SERIALS DATA RECORDING SHEET

PLEASE NOTE (1) When working in the field collect data where possible from inspection of copies and the catalogue of serials where available

(2) Checking from published bibliographies will be done later by the editors, especially for i[illegible] not marked (*)

(3) Do not spend more than 5 - 10 minutes on any problem. Rechecking will be done later by editors. Try and collect at least those marked (*)

NAME OF COLLECTOR

LIBRARY WHERE
DATA COLLECTED

VOLUME (OR ISSUE(S)) OF ITEM FROM WHICH DATA RECORDED ON THIS SHEET REFERS (PARTICULARLY QUESTIONS: (10), (12), (13), (14), (15), (16), (17), (18) and (19), WHERE THE DATA IS VOLUME/ISSUE SPECIFIC

VOLUME ISSUES (Where data does not apply to whole volume) DATE

DATA ELEMENTS THE ITEMS (*) SHOULD BE THE MINIMUM RECORDED

---

*(1) TITLE (In full, as it appears currently)

---

(2) ALTERNATIVE TITLE(S)

---

(3) FORMER TITLE(S)

---

*(4) TITLE IN ENGLISH IF DIFFERENT FROM (1)

---

(5) PUBLISHER

---

(6) PUBLICATION SPONSOR

(i)	Commercial	☐	
(ii)	Learned society or professional body	☐	
(iii)	: (i) + (ii)	☐	
(iv)	Government	☐	
(v)	State monopoly	☐	
(vi)	Educational institution	☐	
(vii)	: (ii) + (vi)	☐	
(viii)	: (i) + (vi)	☐	
(ix)	Private body or firm	☐	
(x)	International organisation	☐	

---

*(7) BEGINNING DATE | *(8) ENDING DATE, IF ANY

(9) COUNTRY OF PUBLICATION

*(10) NUMBER OF ISSUES PER ANNUM

*(11) TYPE OF SERIAL — Periodical ☐ Monographic series ☐

*(12) TYPE OF MATERIAL

Journal	☐	Index to research/theses	☐
Abstracts	☐	Yearbook	☐
Indexes	☐	Fixed period report	☐
Contents list	☐	Conference proceedings	☐
Book review	☐	Legal (legislation, report, articles)	☐
Bibliography	☐	Cases and case notes	☐
Statistics	☐	Accessions list	☐

*(13) NATURE OF CONTENTS

Indicate three main areas by a cross. Tick for minor features

Articles	☐	Statistics	☐
Abstracts	☐	Conference proceedings	☐
Indexes	☐	Cases and case notes	☐
Bibliographies	☐	Accessions lists	☐
Contents lists	☐	News articles	☐
Book reviews and new publications (Not advertisement)	☐	Review articles	☐

*(14) LANGUAGE(S) OF CONTENTS

(15) SUBSCRIPTION PRICE (1969)

(Please indicate where alternative data is used)

(16) COVERAGE BY INDEXING AND ABSTRACTING SERVICES (Where this is listed within the item)

*(17) ABSTRACTS WITH ARTICLES

All main articles ☐ None ☐

*(18) NUMBER OF ARTICLES in 1969
(Listed main article in index/list of contents)

*(19) ASSESSMENT OF SUBJECT CONTENT

Take as guidance the prepared list of subject headings in the Manual and use or modify accordingly

FINAL VERSION

DISISS - SERIALS DATA RECORDING SHEET

PLEASE NOTE (1) When working in the field collect data where possible from inspection of copies and the catalogue of serials where available

(2) Checking from published bibliographies will be done at the editing stage, especially for items not marked (*)

(3) Do not spend more than 5 - 10 minutes on any problem. Rechecking will be done later at the editing stage. Try and collect at least those data elements marked (*)

NAME OF COLLECTOR	LIBRARY WHERE DATA COLLECTED

VOLUME (OR ISSUE(S)) OF ITEM FROM WHICH DATA RECORDED ON THIS SHEET REFERS (PARTICULARLY QUESTIONS: (7), (12), (13), (14), (15), (16), (18), (19) and (20)), WHERE THE DATA IS VOLUME/ISSUE SPECIFIC.

VOLUME	ISSUES (where data does not apply to whole volume)	DATE

DATA ELEMENTS	THE ITEMS (*) SHOULD BE THE MINIMUM RECORDED
* (1) TITLE (In full, as it appears currently)	
* (2) TITLE IN ENGLISH IF DIFFERENT FROM (1)	
(3) ALTERNATIVE TITLE(S)	
(4) PREVIOUS TITLE(S)	
* (5) BEGINNING DATE	* (6) ENDING DATE, IF ANY
* (7) FREQUENCY/ISSUES PER ANNUM (Indicate year taken)	
(8) ISSUING BODY	

(9) TYPE OF ISSUING BODY

Association(s), Society, Prof. Body (Membership instns)	☐	Government (National, Local)	☐	International Organization (i.e., UN, EEC, NATO)	☐
Publisher	☐	Educational institution	☐	Commerical/business enterprise	☐
		Political/pressure groups	☐	Private/individual	☐

OTHERS (Description/type)

---

(10) PUBLISHER (Name)

---

(11) COUNTRY OF PUBLICATION

---

* (12) TYPE OF SERIAL

Periodical ☐ Monographic series ☐

---

* (13) DESCRIPTION OF SERIAL (Tick one category)

Periodical journal	☐	Bibliography	☐	Fixed period report	☐
Abstracts	☐	Statistics	☐	Conference proceedings	☐
Indexes	☐	Index to research/theses	☐	Legal/legislation, report articles	☐
Contents list	☐	Yearbook	☐	Cases and case notes	☐
Book reviews	☐			Accessions list	☐

OTHERS (Indicate type)

---

* (14) NATURE OF CONTENTS (Indicate major categories by cross and minor features by tick)

Articles	☐	Contents lists	☐	Cases and case notes	☐
Abstracts	☐	Book reviews and new publications (not advertisements)	☐	Accessions lists	☐
Indexes	☐			News articles	☐
Bibliographies	☐	Conference proceedings	☐	Review articles	☐

OTHERS (Indicate type)

---

* (15) ABSTRACTS WITH ARTICLES (Tick in boxes)

All main articles ☐ Some ☐ None ☐

---

* (16) LANGUAGE(S) OF CONTENTS

---

* (17) ASSESSMENT OF SUBJECT CONTENT (Take as guidance the prepared list of subject headings)

---

* (18) NUMBER OF ARTICLES IN 1969 (Listed main articles in index/list of contents)	(19) SUBSCRIPTION PRICE (1969) (Please indicate where alternative data is used)

---

(20) COVERAGE BY INDEXING AND ABSTRACTING SERVICES (Where this is listed within the item)